LEVEL II

SDAC-TR-80-2

12

AD A095872

PRESENT STATUS AND DYNAMIC PLANNING FOR AUTOMATIC ASSOCIATION PROGRAMS

John H. Goncz

Seismic Data Analysis Center

Teledyne Geotech, 314 Montgomery Street, Alexandria Virginia 22314

29 February 1980

Sponsored by

The Defense Advanced Research Projects Agency (DARPA)

DARPA Order No. 2551

Monitored By

AFTAC/VSC

312 Montgomery Street, Alexandria, Virginia 22314

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 14 SDAC-TR-80-2	2. GOVT ACCESSION NO. AD-A095872	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 PRESENT STATUS AND DYNAMIC PLANNING FOR AUTOMATIC ASSOCIATION PROGRAMS.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical rept.,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 John H. Goncz		8. CONTRACT OR GRANT NUMBER(s) 15 F33600-79-C-0549, DARPA Order-2551
9. PERFORMING ORGANIZATION NAME AND ADDRESS Teledyne Geotech 314 Montgomery Street Alexandria, Virginia 22302		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS VT/0706
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209		12. REPORT DATE 11 29 Feb ~~ruary 1980~~ 80
		13. NUMBER OF PAGES 47 12 52
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) VELA Seismological Center 312 Montgomery Street Alexandria, Virginia		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automatic Association Earthquake Location

Seismic Signal Processing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This survey reveals that the five functioning automatic association programs existing today in the world, even though they were developed separately in independent institutions, all operate according to the same general schema. Different algorithms and strategies do appear in the event refinement process, and seem to use as much of the attributes of the data as possible to exclude misassociations and false alarms. Recommendations are made for the SDAC AA and ADAPS to use as many of these algorithms as can be supported by the data.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408258

PRESENT STATUS AND DYNAMIC PLANNING
FOR AUTOMATIC ASSOCIATION PROGRAMS

SEISMIC DATA ANALYSIS CENTER REPORT NO.: SDAC-TR-80-2

AFTAC Project Authorization No.:	VELA T/9707/B/ETR
Project Title:	Seismic Data Analysis Center
ARPA Order No.:	2551
Name of Contractor:	TELEDYNE GEOTECH
Contract No.:	F33600-79-C-0549
Date of Contract:	01 October 1979
Amount of Contract	$2,845,057
Contract Expiration Date:	30 September 1980
Project Manager	Robert G. Van Nostrand (703) 836-3882

P. O. Box 334, Alexandria, Virginia 22313

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ABSTRACT

This survey reveals that the five functioning automatic association programs existing today in the world, even though they were developed separately in independent institutions, all operate according to the same general schema. Different algorithms and strategies do appear in the event refinement process, and seem to use as much of the attributes of the data as possible to exclude misassociations and false alarms. Recommendations are made for the SDAC AA and ADAPS to use as many of these algorithms as can be supported by the data.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This document is a comparison and evaluation study of seismological automatic association computer programs. There are five such programs known to be in use, or to have been in use, at the time of this writing. These five programs and the names of the installations using them are:

AUTOMATIC ASSOCIATION OR AA	SDAC
ADAPS	AFTAC
FORAGE_S	NEIS, Denver, Colorado
DEMO	FÖRSVARETS FORKSNINGSANSTALT, Sweden
DETECTION ASSOCIATION PROCESSOR	TEXAS INSTRUMENTS

These five computer programs address the same problem, namely: given a chronologically ordered list of seismic arrival times (plus auxiliary information), produce a listing, or bulletin of epicenters, depths and origin times of earthquakes or explosions, that, in some measurable way, best fits the observed arrivals. Such a bulletin is usually interpreted as a network dependent representation of the true seismicity of the Earth, and confidence in this interpretation increases with an increasing number of stations reporting arrivals for a common event.

When there are only three or four signal arrivals associated to an event, that event is considered somewhat doubtful, its magnitude is usually small, and the validity of the event may have to be verified by a human analyst carefully reviewing the original signal waveforms, improving the time picks and using any other corroborating data available. If it cannot be verified in this manner, or by some other means, the probability is high that the computer association program, no matter how cleverly it has been designed and coded, has produced an unreal event, or false alarm. Certainly one of the major problems here is to keep the incidence of false alarms to as low a level as possible.

With one exception, the five automatic assocation programs discussed in this report have been designed and developed independently from each other at separate installations. The exception is, of course, the close relationship between ADAPS and the SDAC AA; the algorithms and coding of these algorithms for ADAPS which were done by Geotech personnel in the 1960's were transported in as unmodified a form as possible, first to the Network Event Processor (NEP) at the SDAC, and thence to AA. Differences in the coding of ADAPS and AA are due primarily to different ways of accessing the signal arrival list. In ADAPS, a block of signals is held in core during processing; in AA, the signal list is on a disc file and only the few signals immediately being processed are brought into core. The status of these programs at the present time reflects the fact that their development, or expansion, has taken place independently and in an evolutionary manner. The improvements in program operation, in reduction of false events, and in accounting for signal arrivals still unassociated at the termination of processing tend to reflect the distinctive nature of the attributes of the data available at each installation. The difference in the global networks that supply the data to the different installations is also very important in affecting the design and evolution of the five AA programs.

The SDAC AA is the only program whose data stream includes automatic detections from real-time on-line seismic stations. ADAPS and NEIS data streams both consist of analysts' picks and are assumed to be error free. ADAPS seems to contain more auxiliary information, such as signal azimuth and distance and S phase in addition to P. The mission of NEIS is to produce a bulletin of extremely high confidence, which will be distributed to worldwide users, and therefore cannot contain doubtful events. Some signal arrivals must be ignored to achieve this.

In the sections that follow, the five automatic association programs are each presented in narrative form. This method of presentation was chosen so that the overall structure and strategies of each program could be readily grasped. Details of machine dependence, input-output considerations, and system routines have been intentionally not emphasized. The material here is mainly abstracted from technical reports, except for the SDAC AA, for which no comprehensive technical report has been written, and Sweden's DEMO, for which a preprint report exists (a final report is expected to be available in mid 1980).

Following the narrative descriptions of each program there is a section discussing the different algorithms, such as how to find a trial epicenter, with the emphasis here on revealing the strength and weaknesses of the different approaches. This section concludes with recommendations for improving the SDAC AA and ADAPS.

AUTOMATIC ASSOCIATION AT THE SDAC

The automatic association program at the SDAC has its evolutionary origins with the algorithm and coding developed in the 1960's for the ADAPS system. These computer programs were incorporated into the Network Event Processor (NEP), which is an interactive system designed to permit seismic analyst access to a wide range of seismic data, including both waveforms and character information, via a sophisticated CRT display. Indeed, the original design concept for AA considered that its main use would be to quickly associate the signal arrival list, thereby generating a bulletin which the seismic analyst would use as a starting point to refine the event definitions therein, and which would save the analyst the time and drudgery that goes along with the initial phases of event definition.

The section that follows describes the algorithms of AA in essentially the same logical order and flow as occurs during progam execution.

Select Arrival

The first task in automatic association is to get a trial location. The signal arrival list is searched for array beam detections several times from beginning to end to select an arrival according to the following strategy. First, P wave detections are identified by having a beam velocity between 6 and 25 km/sec in order to select a teleseismic signal that is not a core phase. A signal-to-noise gate is used to extract only the larger amplitude detections, based on the strategy that a large teleseismic array detection is probably caused by a large event which would have many reported detections. These detections would be associated and removed from the signal list early in processing, rather than run the risk of falsely associating them with a weaker signal, which could happen if the early passes through the Phase Arrival Queue (PAQ), synonomous to signal arrival list were not restricted only to large amplitude signals. The search to select a beamed arrival takes place on the P-phase, first for large signal amplitude over all reporting beamed stations one by one, then for smaller signal amplitudes for the same stations. Then the search continues from the top for core phase PKPDF, again for large signal amplitude, and then for small signal amplitude.

At the time of this writing, there are seven stations which produce beam information, so the four search stages just described are each carried out station by station. In other words, the PAQ is searched chronologically from beginning to end a total of 28 times. Table I lists the typical search parameters for a run. Since DP was not operating properly for this run, the station cards were withdrawn for the Alaskan net because of the large number of false detections. That is why there are no trial epicenters for BFAK, CNAK, etc in Table I. NB2 is the NORSAR array and it is used first on the assumption that its detection threshold is the lowest of the seven beaming stations. YKA, WRA and GBA do not report a measure of S/N, so their S/N are arbitrarily set equal to 0.0 in the PAQ. The S/N bands in Table I are adjusted as shown so that these three stations will not be queried on the first pass for P, or the first pass for PKPDF, but only on the second pass. Table I shows the number of beamed detections found at each stage for each station; these are processed as described later. Most fail to make an event; the number of successful events made is shown.

On successive passes made through the PAQ, events will be generated and association of signals will be made. Once a signal (beamed or not) has been associated to an event, it is effectively removed from the signal arrival list and cannot be used again. This is also true of array detections that are flagged to be coda signals or side lobes by DP. We expect to investigate the value of changing the permanent association flag to a tentative, or multiple one.

Trial Epicenter

AA relies mainly on the presence of beamed information in the signal arrival list for generating trial epicenters. The arrays are at NORSAR, YKA, GBA and WRA; in addition, there is quasibeam information derived from the six Alaskan stations which results in signal velocity and azimuth information and effectively constitutes a beam. These two pieces of information are processed to yield a location and OT (origin time) on the Earth, and this location is used as the trial epicenter. The initial location may be poor, admittedly, but it is sufficient to gather additional signals for event refinement. One of the problems we have found with this algorithm is that it cannot distinguish between a deep, closer event on the same ray path from a surface event further away, and therefore gives a poor initial location if the actual event is deep (150 km or more). The cure for this problem would be to include depth as a parameter in the event generation process.

TABLE I.

SDAC Automatic Association Search Scheme Used to Select an Array Beamed Arrival

(Velocity 6 to 25 km/sec) P Phase	S/N Band 1	Number of Beams Found	Number of Events Declared
NB2	10 - 999	15	5
BFAK	10 - 999	0	-
CNAK	10 - 999	0	-
TNAK	10 - 999	0	-
YKA	999 - 999	0	-
WRA	999 - 999	0	-
GBA	999 - 999	0	-
(Velocity 6 to 25 km/sec) P Phase	S/N Band 2		
NB2	3.5 - 10	50	6
BFAK	2.0 - 10	0	-
CNAK	2.0 - 10	0	-
TNAK	2.0 - 10	0	-
YKA	0.0 - 999	9	1
WRA	0.0 - 999	28	4
GBA	0.0 - 999	2	-
(Velocity 25 to 9999999 km/sec) PKPDF	S/N Band 1		
NB2	10 - 999	1	0
BFAK	10 - 999	0	0
CNAK	10 - 999	0	0
TNAK	10 - 999	0	0
YKA	999 - 999	0	0
WRA	999 - 999	0	0
GBA	999 - 999	0	0
(Velocity 25 to 9999999 km/sec) PKPDF	S/N Band 2		
NB2	3.5 - 10	9	1
BFAK	2.0 - 10	0	0
TNAK	2.0 - 10	0	0
CNAK	2.0 - 10	0	0
YKA	0.0 - 999	3	0
WRA	0.0 - 999	4	1
GBA	0.0 - 999	0	0

Find Initial Arrivals

Once a trial epicenter location and origin time have been generated, the signal arrival list is scanned sequentially through a time window and each PAQ entry is either accepted as a candidate arrival, or is rejected for one of the reasons given below.

For this scan, the starting time of the window is two minutes before and the end time is twenty-two minutes after the trial origin time, making the window a total of twenty-four minutes. On the average the density of signals and detections in the PAQ is approximately 100 per hour, so this time window contains on the order of 40 signal entries.

As each signal arrival is read, it is loaded into an output array for location refinement (HYPO), unless it is rejected for any one of a number of criteria, such as follows: (at this point a tentative phase is assigned to each signal by comparing its trial travel time with stored Herrin 68 travel times).

- coda or side lobe detection (NORSAR)
- associated to another event
- station is in shadow zone
- this signal has a larger residual than one already accepted into output array for the same station
- phase is not P or PKP
- output array is full (25 entries).

Event Refinement

At this point in the construction of an event, the trial location and origin time, along with up to 25 possibly related arrivals, are passed to a subroutine REFEV, which refines the event.

Iterative calls to HYPO are made and, after each pass, the refinement process consists of deleting arrivals according to the following strategies:

- Delete arrival if in shadow zone
- If standard deviation of residuals is greater than 5 sec, delete all arrivals with residuals greater than 5 sec

- If standard deviation is between 2.5 and 5 sec, delete one arrival with largest residual
- If standard deviation is less than 2.5 sec call find initial again, this time with up to 50 potential signals and repeat above refinement. When enough refinement (deletions) has been carried out, confidence ellipses are computed and the event is ready for declaration.

Event Declaration

After the event refinement procedure just described, the module DCLEVT (Declare Event) is called to perform the following tasks.

- Generate an event number, based on event's origin time
- Store event information and associated signal record numbers
- Update the records of the associated arrivals in the signal list (PAQ); set association flag to 1, store event number, phase code and value of residual.

Since this is the logical end of unit of work in the automatic association program, control naturally goes back to the section which searches for and generates a trial epicenter location and origin time, continuing chronologically until the whole signal arrival list has been examined.

ADAPS: AN AUTOMATIC DATA ASSOCIATION AND PROCESSING SYSTEM FOR SEISMOLOGICAL DATA

The ADAPS is a highly mission-oriented system. The description of the processing flow of ADAPS is presented here so that it can be meaningfully compared and evaluated with respect to the other automatic association programs addressed in this report.

Data reported by seismic observatories, including WWSSN stations, contain station designator, date, arrival time, direction of first motion, impulsiveness of start, phase name, signal type, amplitude, period, estimated direction to source, recording instrument type and recording component. Since all of these data are hand picked by experienced human analysts from analog signals (Develocorder, etc), the data upon which the computer operates are considered error free.

Processing Sequence

Figure 1 is a generalized block diagram of the processing sequence now performed by ADAPS. AFTSHK, designed to quickly associate aftershock data, is a new pass inserted between ASSOC and TRIFX. REVUE, previously the first pass, is now the last and processes all unprocessed unassociated station data recorded on short period vertical instruments instead of only initial arrivals.

The ADAPS system is designed so that data are processed in several passes through the station arrival data file. The data file is processed first by fast techniques which associate much of the data to larger events before more time consuming techniques are applied to harder-to-associate data from weak events.

ASSOC

The first pass through the data file (ASSOC) is made to associate new arrival data to events input manually and/or that have been determined on previous runs without all possible data having been received. When any new data are associated, event parameters and classification values are recomputed

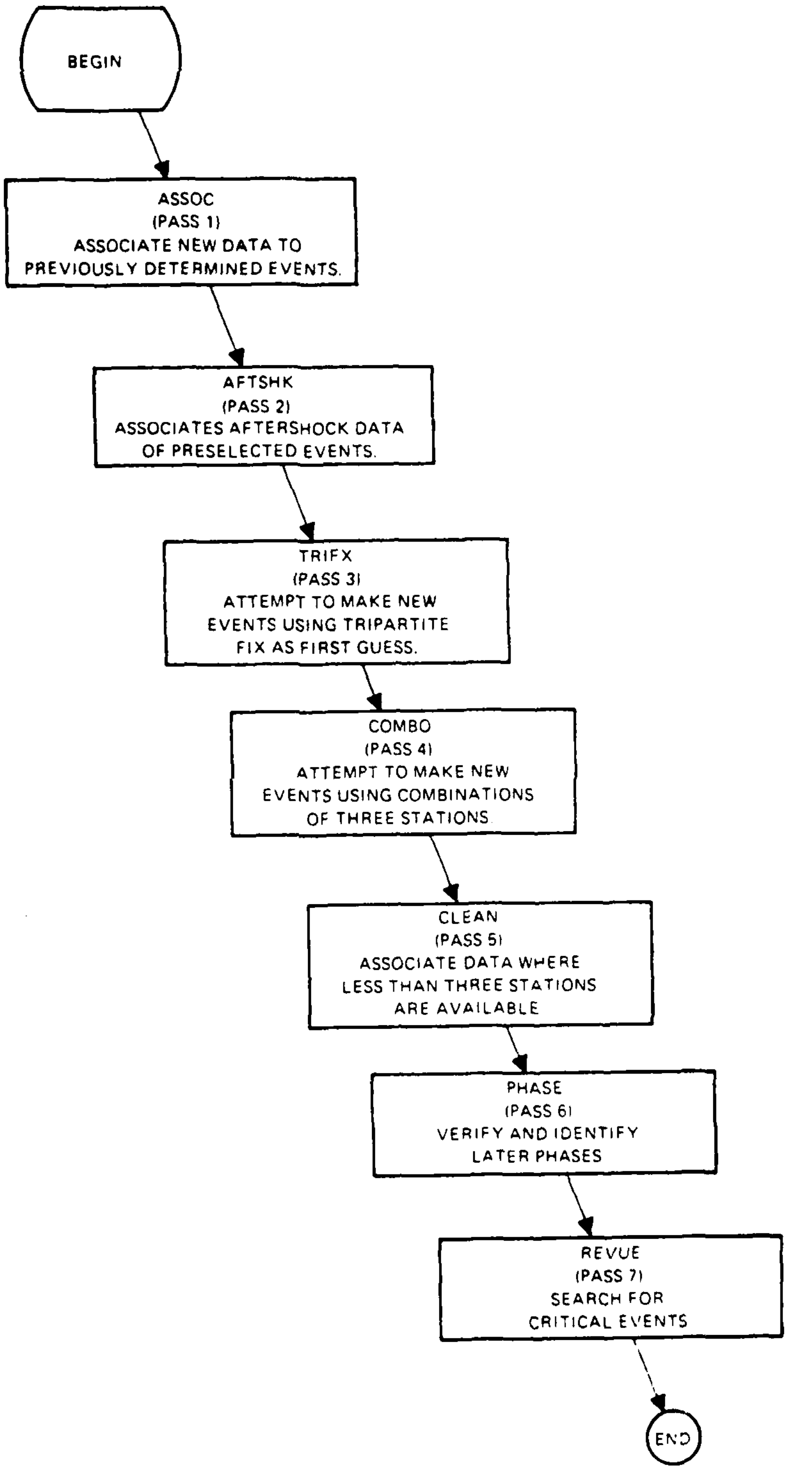

Figure 1. ADAPS processing sequence.

and processing control data of the event and of each arrival used in the solution are updated. The procedure of looking for new data to associate to previously determined events is continued until the store event files (final and working) or data file have been exhausted.

The processing logic of ASSOC has been modified to provide the capability of processing events under various options denoted externally by an analyst. In the ASSOC pass, the final event file is searched for events externally designated for special processing. These events might include those needing complete updating because they were added to the file manually by an analyst or those needing relocation or magnitude recomputation because the analyst changed the associated data.

After the required special processing is completed, unprocessed unassociated station data are checked for association with the events in the final event file and, then, with working events. When any new data are associated, event parameters and classification values are recomputed, the processing status of each arrival used in the solution is updated, and the event processing status is set to indicate that the event has received further processing.

ASSOC contains the necessary logic to prevent any event from being processed more than once during any one execution of the pass unless previously associated data are released from association when a hypocenter is deleted from the file. Processing with ASSOC continues until each new arrival has been checked for association with each final and working event (in the acceptable time frame) or when all final and working events have been processed.

The AFTSHK pass is intended to reduce processing time during periods of high data volume caused by aftershock series or earthquake swarms. The method used in AFTSHK is primarily one of pattern recognition; the station data file is searched for arrival time sequences that correspond to the pattern produced by the main shock. The routine has provisions to search for and process five different patterns.

Station data used for the main shock are stored in increasing order of travel time so their sequence is the expected arrival time sequence for after shocks. The data file is searched for an unprocessed unassociated initial arrival at one of the stations stored in the aftershock array. When one is found, it is designated the base station and an expected time of arrival is computed for each of the stations which, if this first arrival is from the same general location as the main shock, should have arrivals later. The expected times are computed by adding the travel time differences between later stations and the base station to the observed arrival time at the base station. A tolerance of ± 4 seconds is allowed to account for slight hypocenter variations and timing discrepancies. If the expected station reports an unassociated initial arrival in the correct time frame it is assumed as a possible association to an aftershock.

The process is repeated for all stations in the aftershock array with expected times later than the base station. After all have been searched for and if at least four have been found, AFTSHK calls routine HYPO to refine the epicenter estimate and routines AQUPS to add the event to the final event file. To increase the speed of processing when the data volume is extremely high, a control parameter KWIK controls the degree of refinement attempted on the aftershock events. This parameter is designated by the analyst before running ADAPS and indicates either that additional data are to be associated, but not used to refine the event parameters, or that the event is to be completely worked by routine FATE and all possible data associated.

AFTSHK is not completely automated; locations suspected of having aftershock activity must be input by an analyst who, before ADAPS is run, notices a high amount of activity in the data file and identifies the source of the activity.

TRIFX

The third pass through the data file (TRIFX) is the first attempt to estimate the location of new events. The pass is started after all the events in the current event list have been checked for associated arrival data and the station data file still contains unprocessed unassociated arrivals. The initial epicenter estimate is determined from a combination of three initial arrivals at the stations E, IM, TT, and BC. These arrivals are processed with a fast tripartite location routine (QFIX) which, under the assumption that the distance between stations is small compared to the distance to the epicenter, estimates the distance and azimuth to the event by computing the apparent surface velocity across the tripartite and the azimuth of approach. This epicenter estimate (working event) is then used to check for other associated data and, if confirming station data are available, the hypocenter is refined and the event is moved to the final event file. This process is repeated until all the unprocessed tripartite arrivals have been examined.

COMBO

After all possible tripartite fixes have been determined, pass 4 (COMBO) attempts to produce new epicenter estimates from logical combinations of three stations with unassociated initial arrivals. The combination of arrivals likely to yield a correct solution is selected by comparing reported event types and observed arrival time differences to the distances between stations. Epicenter estimates are computed by using the three-station location routine (TRIX) and the most likely TRIX solution (several solutions may be obtained from one combination of arrival times) is selected by checking event type agreements, and reported directions. Other unassociated arrivals are checked for association first to the location logically selected as most likely and if no associated arrivals are found, the next solution is checked, etc. If no other data appear to assoicate with any of the TRIX solutions, these solutions are retained in a working event file and processing is continued.

CLEAN

Pass 5 (CLEAN) is the final attempt to associate remaining unassociated initial arrivals by computing a new event location. Because events with at least three stations reporting an initial arrival have been determined in previous passes, this pass uses other data reported by the station in addition to its initial arrival time to produce epicenter estimates. One-station epicenter locations are produced when a station reports the initial P phase arrival time, identifies the S phase, and reports the direction to the source. Two-station fixes can sometimes be made when one station identifies the S phase and the other reports a direction. These epicenters are not accurate, but in many cases will be adequate to eliminate from further association some data which could be erroneously associated to some other event or remain unexplained in the data file until reviewed by an analyst.

PHASE

PHASE, the sixth pass of ADAPS, attempts to verify later phases called by station analysts and to identify any reported unidentified arrivals as phases to an established event. In addition, an attempt is made to identify unassociated initial arrivals as possible PKKP or P'P' phases when the event is of magnitude greater than 4.8.

PHASE was modified to reprocess an old event (one not made or reworked on current ADAPS run) only when an unprocessed arrival is within 40 minutes of the origin time. Routines SURF and SURTT have been added to include surface wave verification. PHASE was also modified to not set association status of an identified phase unless the station analyst associated it with an initial arrival associated to the event.

Processing is begun by finding the first final event unprocessed by PHASE during the current ADAPS run. The portion of the station data file with arrival times later than the event origin time and earlier than 40 minutes after the origin time is searched for an initial arrival, either unassociated or associated to the event, and from a station not previously processed by PHASE with this event. If the current event was not newly made or reworked during the run of ADAPS the arrival must also be unprocessed. All indices of arrivals from this station which are later than the arrival discussed above but earlier than 40 minutes after the event origin time and are either unassociated or associated to this event (and unprocessed if the event is old) are stored in a working array.

The data in this array, including epicentral distance, recording component, instrument type, and observed travel-time are checked to determine the nature of the reported arrival. These checks are performed to eliminate erroneous identifications due to similar travel-times of different phase types and to reduce the number of phase identification candidates. Checking distance separates P and PKP distance range phases, component separates vertical and horizontal phase types and travel-time separates direct, singularly-reflected, and multiply-reflected phases. Results of these checks determine the appropriate subroutine necessary to identify the arrival. Within each subroutine, distance and observed travel-time is checked again to further reduce the number of phase candidates.

When a secondary arrival associated to the event being processed has been identified by the station analyst, an immediate verification attempt is made; however, if the phase is not verified, the program makes an attempt to identify it, although the computed identification will not override the analyst's designation.

For both identification and verification procedures, the residual (observed travel-time minus computed travel-time) is compared against residual time windows of each phase candidate. If the residual falls within one window, identification or verification is complete; if two or more candidates exist, the one with the smallest residual is chosen to identify the arrival. When residuals fail to fall within any of the windows computed by the particular subroutine used, no identification is made.

REVUE

The last process control program is REVUE. This process scans the data file for arrivals which may be associated with an event at a critical location. Processing control options define data eligible for correlation. The geographical locations of several known or suspected test sites may be stored and the reported data are checked for correlated time differences. If three or more arrivals meet the acceptable criteria, the data are output.

REVUE originally was the first pass in the processing sequence because the design was based on the ADAPS system being linked to the remote display terminal so that immediately after correlation was found, association of additional data, hypocenter determination, and magnitude computation would be performed and the results displayed while the system then continued to process the routine data. Under the present configuration of the 360/75I computer system, ADAPS is not able to access the display twice during processing so the analyst will receive all results only at the termination of all ADAPS processing. This changed the design concept of early detection so REVUE was made the last pass and instead of only processing initial arrivals (P-waves), it now processes all unprocessed, unassociated arrivals which were recorded on the short-period vertical instruments.

Use of Arrays and Beamed Information in ADAPS

Because ADAPS was developed and put into use before array information was available and reliable, it is not structured to use azimuth and velocity information to routinely find a trial epicenter, as does Geotech's AA.

Gerry Clawson offered the following comments in response to the question of how ADAPS can use array azimuth and velocity, or arrays in general.

- For the giant arrays like LASA and NORSAR, 3 or 4 individual seismometers could be designated for input directly to HYPO, or could be considered components of a tripartite array for input to QFIX.

- In the CLEAN pass there is provision for using station location, direction and distance to source to associate any unassociated signals by computing a new event location. This information could be derived from azimuth and velocity, as well as S-P times, as is done in CLEAN.

THE SWEDISH AUTOMATIC ASSOCIATION PROGRAM - DEMO

Försvarets Forskningsanstalt (Defense Research Institute), Sweden

Sweden has offered to establish and operate an International Seismological Data Center as part of an international global monitoring system for the verification of a Comprehensive Test Ban Treaty. The automatic location and association program (DEMO) discussed here was part of a demonstration conducted at Hagfors Observatory in July 1979. The information about this program is derived from a preprint report (Dahlman and Sidor, 1979) and from a personal visit by Dr. Robert R. Blandford, Teledyne Geotech, to Förvarets Forskningsanstalt.

DEMO used array information from Yellowknife in Canada, NORSAR in Norway, Grafenburg in Germany and Hagfors in Sweden. The correct association and the definition of new events is based on the estimated azimuth and distance from the array station to the event. More than 70% of the events defined and located by DEMO, but not reported by USGS, were based on data including preliminary location data from at least one array station.

The following section is Dr. Blandford's memo dated 11 February 1980 concerning the information he obtained by visiting Försvarets Forskningsanstalt.

"The Swedes seem to have a nice data management structure and a very nice AA for putting together arrival times (the kinematic algorithms) from which we can pick up some useful ideas. They also have a method of throwing out arrivals with inconsistent amplitudes (the dynamic algorithms) developed by Eva Elvers. The idea is good but in application it seems only to hurt. Elvers thinks that even from a theoretical point of view it doesn't work and she thinks that a lot more work is needed on it.

I will first discuss the kinematics algorithms in some detail, then the dynamic, and finally a few words about the data files.

The kinematic algorithm is the work of Raynar Slunga. The data consist of P and S arrival times and azimuth and slowness from the usual arrays plus Hagfors. In addition to NEIS data and MOS (Moscow) data they reviewed all the SRO SP detections, using their PDP-15, and merged these times and amplitudes into their data base for the one week comprising the study, January 15 to 21, 1978.

The AA algorithm begins by predicting arrival times from array locations. Phases which fit are put into a location algorithm and an initial location is obtained. If an arrival has been designated as 'local' by the reporting station and if the station turns out to be more than 15° from the epicenter, then that arrival is deleted. If the P-S distance can be available, and if that distance differs from the station-epicenter distance by more than 1° (I am not sure of this parameter) then that P arrival is thrown out. If an arrival passing both these tests is in error by more than about 25 seconds (the actual value depends on the size of the confidence ellipse using an algorithm which I do not know) then that arrival is thrown out. PKP arrivals are associated to the event but not used in location. All arrivals used in location are removed from consideration for future events. Locating arrivals are called defining arrivals. Array arrivals will also be removed if the measured azimuth and slowness do not agree with those calculated from event and station location. An event will be defined and put out if 4 parameters fit, e.g., 4 arrival times or 1 arrival time and 3 measurements at an array.

After all the array trial epcienters have been considered in time order the a 'combo' like procedure is begun. Slunga has in effect reinvented COMBO and has a nice way of sweeping through the queue, which is the same as used in ADAPS.

The processing for each triplet is to find two solutions at four fixed depths each. Depths might be 0, 100, 200, 400 km. The solutions are found in some way by trial and error, I believe Slunga starts with a trial solution at the station with the earliest arrival time. How he moves around to find the solution from this point I don't know. After one epicenter has been found then he takes the initial guess for the second event on the other side of the earth. Each of these 8 epicenters is tried until one of them generates a final event by the method discussed above for events started using arrays. (We would be able to use TRIX for this purpose simply by calling TRIX with an index to tell it which depth travel-time curve to use.)

The dynamic aid to AA first computes the magnitude using Ringdal's maximum likelihood formula; allowing the magnitude to be influenced, for small events, by up to 40 non-detecting stations. An interesting feature of this calculation is that there are four categories of stations, I. defining arrivals with amplitudes; II. defining arrivals without amplitudes...for this case the a-priori noise level is taken as a lower bound to the signal amplitudes; III. associated signals which, however, have too large a residual to be called defining...in this case the noise...or signal level if available...is taken as a lower bound to the signal amplitude; IV. no detection, and no data...in this case the a-priori noise is taken as an upper bound to the signal.

In as much as possible, station down-times are kept track of and if the station was down it does not figure in the magnitude calculations.

After the magnitude is calculated, an assessment is made of how 'likely' the particular situation is; and if it is not very likely, then the individual station with the least likely occurrence (e.g. a very large signal or no detection even though 1° from the event) is not allowed to contribute to the event and the process is repeated. What happens in practice is that often all the stations with observed amplitudes get thrown out. Elvers thinks that she has not figured out how to evaluate how likely the particularly situation is; and so the algorithm throws too much away.

The data file structure was specially developed by Gunnel Barkeby and a consultant. Gunnel visited the SDAC 5 years ago she says. They felt that 'canned' DBMS could not handle the calculations with which they wanted to surround their data. Note that all of this system resides on a 370/165; and 7 days of AA runs in 40 minutes. From the data base it is possible to select a data time window; to list out all arrivals to any station together with the events to which it has been associated or used to determine; to list out all events with arrivals; to list out down times, etc, the access to the fields seems very complete. It is possible to select an event, delete arrivals and relocate. In the relocation it is possible to change all parameters of the location program e.g. time, azimuth and slowness residual tolerances, travel time table used e.g. Herrin or JB, distance-amplitude formula used in the dynamic part of the program e.g. Veith-Clawson or Gutenberg, they use Veith-Clawson routinely. In general, you have good control over the situation. The system could not, however, be easily used at present for bulletin preparation by an analyst because of the lack of an FI command. I mentioned this to Gunnel Barkeby and she said that she might try to put that in.

Their bulletin has not been evaluated in a critical way as we have done for our AA; for example, they have the same event three times and have not 'fixed' the program so that that does not happen. I tried to impress on them that there is reality out there and that we thought our bulletin 'looked' good too until we compared it to what Boomer got. Probably their bulletin is pretty good though since nothing but human picks went into it. I have come back with the full bulletin of events, plus the supporting arrivals for the events on the 18th through 21st and it might be worthwhile for Boomer to critique it to the extent that is possible from the data we have from them for the 18th-21st, together with our own data sources for that time period."

FORAGE_S: THE NATIONAL EARTHQUAKE INFORMATION PROGRAM - U.S.G.S.

The early work on an automatic location and association computer program is described in a paper by Engdahl and Gunst, 1966. It uses five (plus an optional sixth) arrivals (contained within an 11 1/2 minute time window) to determine a trial epicenter. It then follows a fairly straightforward way of finding possible associations and refining the hypocenter by excluding those signals whose residuals are above certain criteria.

The original program, called COAST, has gone through evolutionary stages and exists at the time of this writing in a PL-I language program named "FORAGE_S." Dr. Gerald Dunphy, the scientist responsible for FORAGE_S, says he has followed the basic strategy of Engdahl and Gunst.

The following pages giving the narrative description of COAST are abstracted from the Engdahl and Gunst, 1966 paper.

"GENERAL SYSTEM DESCRIPTION"

"A punched card file, in which the P arrival times are arranged in chronological order, forms the source of program input on magnetic tape. This file is updated before each computer run by adding:

(1) Newly arrived data keypunched from preliminary station bulletins;
(2) Data cards obtained by computer processing of telegraphic communications from observatories;
(3) Control cards which resulted from processing the output of the previous execution of the COAST Program. These control cards contain the coordinates of either unpublished hypocenters which are to be recomputed or hypocenters which are acceptable for publication.

Data are processed by the program in blocks of 300-500 P arrivals and associated phases at one time. The exact number is determined by the program and depends upon the number of observations which could possibly be within range of the origin times supplied by the control cards. Each block of data is processed in three steps (See Figure 2). The first step in processing is to scan the data block to eliminate readings consistent with already published results. Next, the control card coordinates of previously determined but unpublished positions are used as first approximations, and refined hypocenters are recomputed with all associated data now available. Finally, the residue of data is systematically searched to obtain groups of five or more compatible time readings. Each group is processed separately by first using five selected stations to locate an approximate hypocenter and then using all the available compatible data to determine a refined hypocenter. Whenever the cycle is completed, a new data block is read in.

Upon termination of the COAST run, a list of the output tape is examined by seismologists and entries made on an output allocation sheet for each refined hypocenter. These entries are transferred to punched cards which govern a programmed processing of the output tape and provide for punching of control cards, listings for publication and other editing needs. The control cards carry, in addition to hypocenter coordinates, information which the seismologist wishes to communicate to the computer, such as a request for depth restraint, or for removal of stations from the solution. These will be performed automatically on the next computer run.

DETERMINATION OF FIRST APPROXIMATION

The determination of first approximation to previously unlocated hypocenters from a data file known initially only to be in chronological order by P arrival time poses a formidable problem. A large part of the input data file is made up of related data which are insufficient for accurate locations, related data overlapped with one or more other earthquakes in time, data totally unrelated to other P arrivals, and PKP arrivals erroneously identified as P.

To resolve these problems it was decided to proceed sequentially through the data file until a group of readings could be found which fell within some specified time range, and then to make the best use of the little information which is available (time order, phase identification, S wave arrival time and station rating) to determine if the reported arrival times for preferred data within this group were possibly related (see Figure 3). It was found by trial and error that a time group such that the difference in time between the earliest and latest station readings is not greater than 11 1/2 minutes, offers the best probability of not also including PKP readings from the same earthquake. After such a time group has been found, the observations within the group are examined in an effort to select from the group the five stations most likely to produce an acceptable provisional hypocenter. This is accomplished by eliminating stations on the basis of the compatibility of S arrivals or of P arrivals alone when S is not reported. Preference is given to prime data which include those stations reporting an S arrival and/or rated as prime stations with the requirements that the P arrival be well identified (P, eP, iP). Prime stations have been designated on the basis of a subjective judgment of the reliability of reporting, instrument sensitivity and/or critical geographic location.

As a simple illustration of station compatibility criteria, consider the case of two stations not more than 100° apart, which report P arrivals but no S. If the difference between the P arrival times at the two stations is greater than the travel time for the distance between the two stations, the readings are incompatible.

Upon completion of the reduction of the original time group to compatible prime data we choose the earliest arrival and five other stations, one from each of five evenly spaced time intervals within the reduced prime data group. The arrival chosen from the last time interval, the most distant station, is designated the optional station. This method of selection not only provides a distance distribution but in most cases has shown to afford a fairly uniform

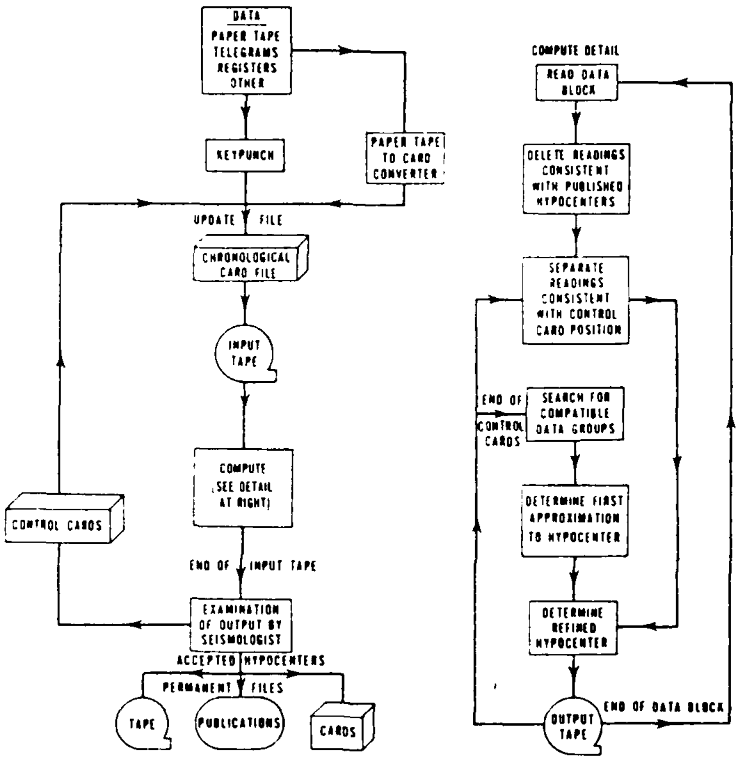

Figure 2. General flow diagram for COAST hypocenter determination program (from Engdahl and Gunst, 1966).

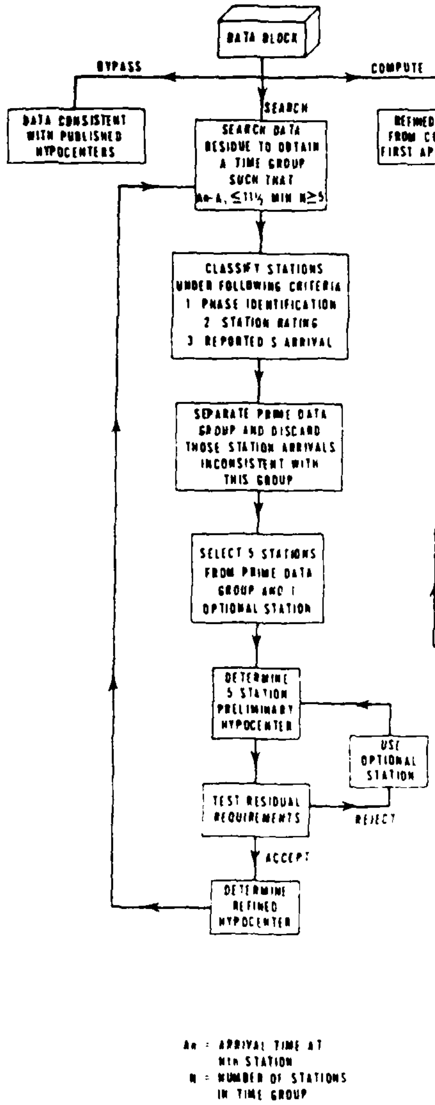

Fig 3.

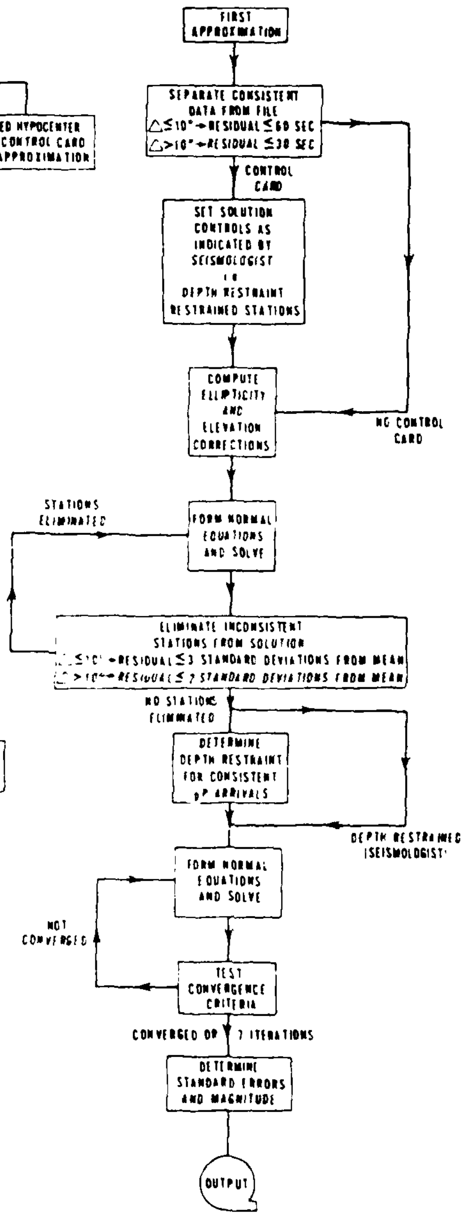

Fig. 4

Subprogram for determination of the first approximation to the hypocenter (from Engdahl and Gunst, 1968).
Subprogram for determination of the refined hypocenter (from Engdahl and Gunst, 1968).

azimuth distribution. Using the five selected arrivals other than the optional station, a hypocenter determination is attempted by the least squares iterative method previously described. In the event the iterative process does not converge to a solution or converges so that the sum of the absolute values of the residuals for the five stations is greater than 30 seconds, the optional station is substituted successively for each of the original five until a satisfactory solution is obtained. Provision is made for the case where there are only five stations in the original time group and thus no optional station. With a successful computation of a first approximation, all other relevant data are separated from the file and a refined hypocenter computed using all station arrivals.

DETERMINATION OF REFINED HYPOCENTER

After a first approximation has been obtained from either a control card or the data residue file, the consistent data must be separated from the data file for computation of the refined hypocenter (see Figure 4). Beginning with an arrival not more than 60 seconds earlier than the provisional origin time, each data point (arrival time) is examined in turn under the following rejection criteria: if the distance is less than or equal to 10°, the allowable residual range is ± 60 seconds; if it is greater than 10°, ± 30 seconds. If the first approximation was obtained from a control card, the various modes of computation and station restraints that were designated by the seismologist are implemented at this time. Next, the P arrival times are corrected for ellipticity and station elevation. The least squares normal equations are now formed and solved with stations eliminated according to the following criteria: For distances less than or equal to 10°, the allowable residual range is ± 3 standard deviations from the arithmetic mean of residuals; if it is greater than 10°, ± 2 standard deviations from the mean. This elimination process is continued by iteration until no further stations are eliminated from the solution. At this point continued iteration is temporarily terminated (assuming the depth has not been restrained by control card) so that depth phases reported by stations remaining in the solution may be examined for consistency. In the event that at least two pP arrivals can be found such that the depth as indicated by each arrival does not deviate by more than 10 kilometers from the mean depth, then the depth will be restrained at this mean level for the remainder of the refinement. On the other hand, if there were not any pP readings reported or they were inconsistent under the above criterion, the depth is determined by the P arrival alone. After the elimination of the larger residuals, examination of the depth phases, and setting of the proper controls, the program enters the final iteration sequence using all remaining stations. Convergence or termination of this sequence is obtained when the absolute value of the arithmetic mean of the residuals is less than or equal to 0.05 seconds or when seven iterations are completed. In the event a negative depth is obtained in any of these final iterations, the depth is automatically restrained at 33 kilometers for the remainder of the refinement. Normal operation indicates one to two iterations for convergence for most determinations.

Unlike the computer program described by Bolt (1960), no attempt is made to attach weights to the residuals.

Following convergence, the standard errors are determined from elements of the inverse of the normal equation matrix and the magnitude (m_b) is computed. All original input information, phase arrivals, periods, amplitudes, magnitudes, distances, azimuths, residuals, hypocenter coordinates and standard errors, are output on magnetic tape. Use is made of the STRETCH disk file to also include a scheme for geographic identification devised by Flinn and Engdahl (1965) which provides a geographic designator for each square degree of the earth's surface.

GENERAL CONSIDERATIONS

The great volume of data processed at the U.S. Coast and Geodetic Survey necessitated that COAST be specifically designed for speed and efficiency. Using the STRETCH computer, a normal file of 30,000 P arrivals on magnetic tape as input can be completely processed in 15-20 minutes with an average output of 200-300 refined hypocenters. Nearly all this time is used in input-output operation since a single iteration of a 100-station hypocenter takes less than 0.5 second. A maximum of 500 stations may be used in any single refinement.

Flexibility was a major consideration in the programming. By means of the control cards the seismologist can exercise the following optional restraints: (a) hold the depth fixed, (b) hold the epicenter position fixed and compute depth and origin time, (c) hold all parameters fixed and compute station time residuals, distance, magnitudes and azimuths. He may also override station rejection criteria by insisting that certain data be retained or rejected from either or both the hypocenter and magnitude computations.

In addition to the routine mode of operation, the entire tape file may be processed in several optional modes. By means of a program-header card computations may be limited to either previously determined but unpublished hypocenters, published hypocenters, undetermined hypocenters or any combination of these options. Processing may further be restricted to a prescribed geographic area and/or a minimum number of reporting stations.

The program can be used to compute distances, azimuths, expected arrival times, or travel times for any hypocenter to either selected stations, or the whole station file of approximately 1,500 locations.

ADDENDUM

After nearly two years of operation without substantial change, this program is currently being revised so as to include several features which will improve the accuracy of hypocenter determination. These will include: (1) The incorporation of new travel time tables being developed from explosion and earthquake data. (2) The application of station time corrections determined as as part of the above study. (3) The ability to employ local travel time tables in those areas where they have been determined. (4) Weighting of observations according to the confidence limits of the new tables. (5) Expanded error statements. (6) The use of $\partial T/\partial \Delta$ and $\partial T/\partial h$ tables instead of polynomial interpolation of the travel time tables."

DETECTION ASSOCIATION PROCESSOR - TEXAS INSTRUMENTS

The detection association processor carries out the function of transforming signal detections obtained independently by a network of seismic stations to seismic event detections. The location, depth, and origin time of the estimated seismic event focus are then used to edit waveform segments associated with the events by estimating arrival times of the seismic phases, particularly the P-phase.

This computer association program, reported by Snell in 1978, was developed in a laboratory environment and has not yet been tested with real data. It is designed to handle multi-array information, single station information or a mixture of the two types of received data. In the multi-array mode, the trial location (for initial association) is found by processing location and distances of multi-array station information (up to five) by a Kalman filter whose output is single, optional combination of the input locations.

In the single station mode, a novel shrinking grid is used to obtain the best trial epicenter consistent with a cluster of signal arrivals. The shrinking grid covers depth and locations for possible P phases up to the shadow zone, and the author believes it can be used with other phases, too, such as PKP.

In telephone conversations with the author, now at Lincoln Laboratory, he stated that he believed the shrinking grid idea was good for the following reasons:

- It is extremely fast, and the iteration loops can be coded tightly.
- The amount of computation required by the grid goes up only linearly with the number of reporting stations, whereas in a combinatorial approach (e.g. ADAPS's COMBO) the amount of computation goes up faster than linearly.
- In any given cluster of signals, the useless ones are thrown away immediately and do not have to be reconsidered.

The following sections are extracted from Snell's report and describe the association modules. Subsequent processing after these stages is standard, and uses HYPO to refine the event location.

The reader is referred to Snell's report for a full discussion of the simulation evaluation of the detection association processor. Basically, the system is shown to work as designed, and several areas where improvements could be made, particularly in interfacing between algorithms, are suggested.

Associator Modules

There are three major associator modules. These are the Array Network Processor, Single-Site Network Processor, and a Mixed Network Processor. The Array Network Processor operates on detection bulletins from automatic detectors at array stations. The Single-Site Network Processor operates on detection bulletins from automatic detectors at sites providing only a single vertical component measurement of signals. The Mixed Network Processor operates on detections from a network of both types of stations. The subsystem functions performed by the network are as follows:

- Real-Time Bulletin Processing
  - Add new bulletins to the current stack of detection bulletins in the detection bulletin buffer
  - Maintain ranking of the detection bulletins by Z-statistic
  - Remove aged bulletins from workspace after a fixed number of failures to associate
  - Remove bulletins from workspace which are associated with previously located events.
- Association Processing
  - Set up one or more preliminary trial locations
  - Refine the trial locations and repeat the association tests
  - Iterate the association tests until a specified number of detections are associated, e.g., four detections, and a locatable event is declared.

- Processing
  - Perform final location of a declared locatable event with arrival times from associated detection bulletins
  - Use depth grid and depth constrained locations to find best estimate of the depth.

Figure 5 illustrates the flow of decisions and actions performed by the detection association processor. The inputs are detection bulletins from seismic stations, either arrays serviced by an array module or single-site stations serviced by a single-site module. The detection association processor operates on a fully loaded buffer of the most recently acquired detection bulletins sent by seismic station detection processors. As space becomes available in the bulletin buffer, the detection bulletin buffer is fed new bulletins from a queue. This space becmones available in several ways. One way is if a bulletin fails to associate with different events some specified number of times, e.g., three times. This aging process is a critical factor in managing the workspace of the detection bulletin buffer. If the aging process is not carefully designed, bulletins will age out because the workspace is clogged with coda detections from preceding larger events. This is definitely a problem with the DAP models described in this report and this problem needs to be attended to if better results are to be obtained. Another way for a detection bulletin to be discarded is for it to have been associated with an event already lcoated and registered on the event list generated as an output of the system. Upon discarding a bulletin in one of these ways, a new bulletin is fed in serially from the queue and inserted into the detection bulletin buffer. The insertion is rank ordered by the magnitude of the detector output, so that the most probable event detections are further up toward the top of the stack. First, all of the array detections are tested for association with the presumed location of the top ranked detection. If the association tests are passed, the location is updated by the Kalman location algorithm. Otherwise, a counter is imcremented to register the number of times the detection bulletin failed to associate.

After completing the association tests of all of the array stations, tests are performed on the single-site detection bulletins. This is done by centering a grid of potential locations on the location estimate obtained by

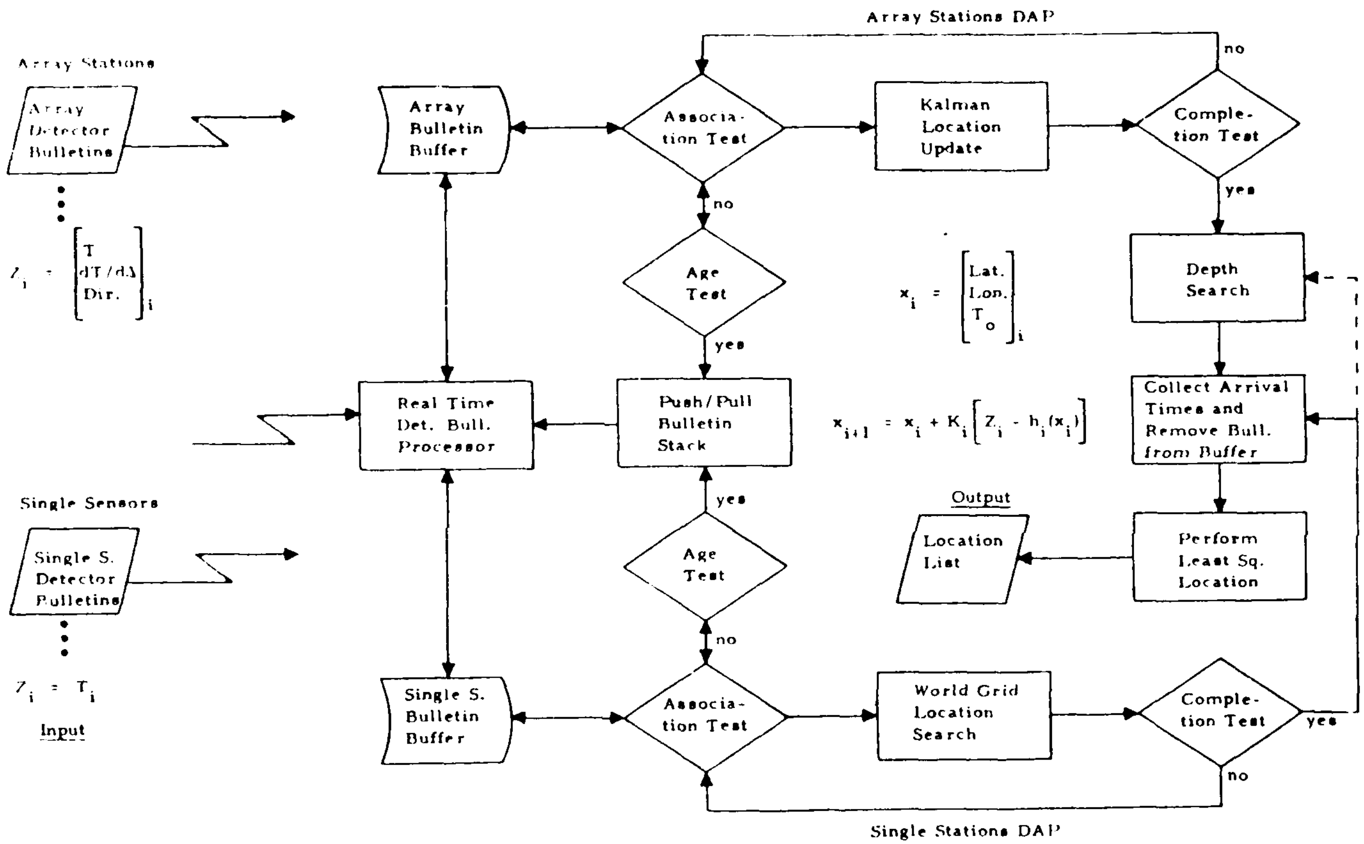

Figure 5. Automated network location (from Snell, 1978).

the preceding operations of the array module. If there are no array detection bulletins, then the location of the station with the largest detector value is taken as an initial location, about which a grid of potential locations is centered. See Figure 6. Association is performed by the single-site module by searching for the maximum number of origin times consistent with a depth location at each grid point corresponding to a trial location. The origin time is said to be consistent if the difference between the origin time of a candidate detection bulletin and that of the top ranked 'key' detection bulletin is less than the expected timing error implied by the distance between the grid points of potential event locations. To improve the precision of the location search, the single-site module association tests are repeated, in toto, several time. At each stage, the output location estimate is taken as a new initial location. The grid size is halved and the search is repeated to find the maximum number of associates with consistent origin times. This is repeated until a prescribed precision is attained for the depth constrained location estimate. After completing the single-site associatior tests, a depth search is performed to find the focal depth with minimum variance error in predicted station arrival times. This search is also performed successively on a grid of depths. The search is carried out to the desired level of precision by successive binary partitions of the grid distance between the preceding best depth estimate and neighboring grid points. Finally, all of the arrival times of associated detection bulletins are used for a linearized least squares location estimate.

The rationale for using the Kalman Filter or collapsing grid method to perform the DAP is to develop a general systematic procedure for selecting detections associated with an event out of a detection bulletin list consisting mostly of false alarms from noise, coda of preceding events, and later phases of preceding events. The main difficulty is to be able to associate and locate smaller events shadowed by much larger events occurring about the same time. The Kalman Filter provides a mathematical basis for obtaining sequential estimates of location and location errors. These can be

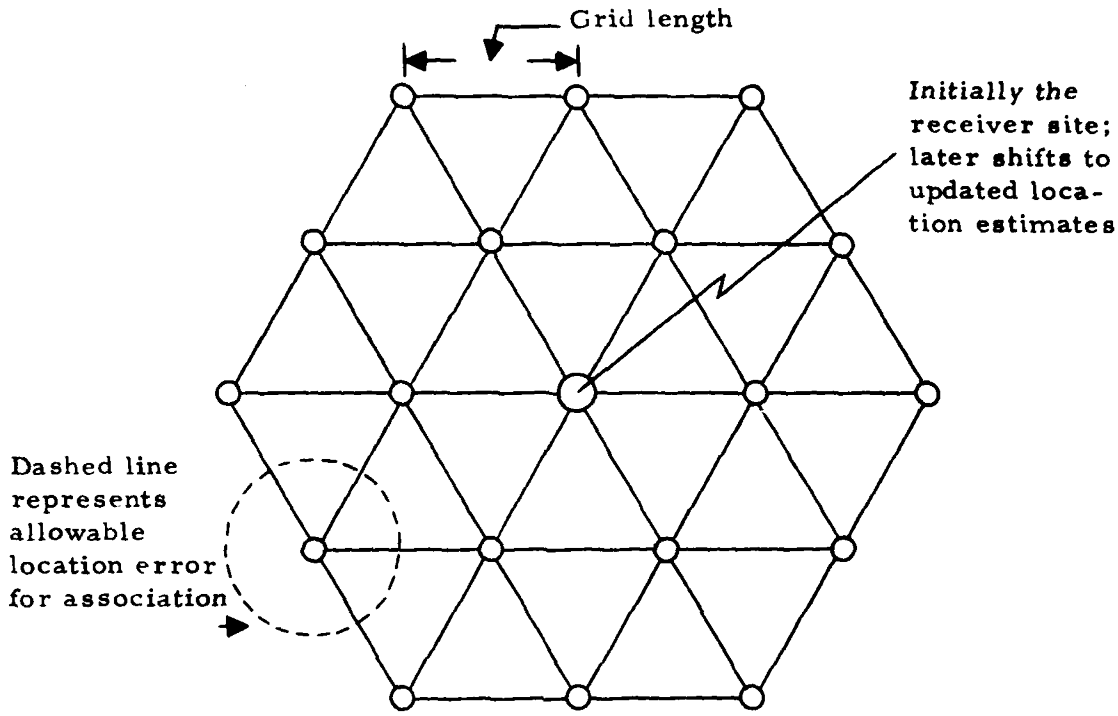

Figure 6. Collapsing grid location strategy (from Snell, 1978).

used as a basis of association testing. The collapsing grid can be used to efficiently weed out false alarms. These inadvertently get into the associated set of detections, causing false event declarations and also causing the possible loss of bulletins which otherwise would have been correctly associated. The algorithms used to perform the previously described subsystem functions still need to be improved for them to operate satisfactorily at high false alarm rates mostly caused by event interference.

Comment

It is not apparent that the methods used in DAP could significantly improve the performance of AA, if they were incorporated into the latter program. One of the difficulties in assessing the significance of DAP is its inability to use available datasets to produce results that are readily comparable with AA. Perhaps it can best be viewed as an experimental program which requires more development and evaluation effort, and this should be supported as permitted by available resources and priorities.

CRITIQUE AND COMMENTS

The preceeding sections have shown that, even though these five automatic association programs have been independently developed in separate institutions, the programs' basic strategies of operation are remarkably similar. So similar are they, that it seems possible to represent all five programs by a single elemental flow chart, as shown in Figure AA, with the second order differences between them shown as comments to the principal algorithms. This scheme shows that the differences among these automatic association programs exist in the way they get started, i.e., in the technique used to obtain a trial epicenter, and in the algorithms used to exclude non-associatable signals,

Besides the shrinking grid method of DAP, there are three techniques for getting a trial epicenter: array beam information, three-way compatible station combinations, and five-way combinations. It is an essential question to ask if one of these three methods is superior to the others.

Certainly the answer must in part depend on the attributes of the signal arrival list. Analyst picked signals comprise the list for NEIS and ADAPS; there is no beam information in the data set, therefore the array method cannot be used with these two datasets. The use of beam information is conceptually simpler and probably computationally faster than searching for compatible station combinations, either three or five. The success of an array-beam-driven automatic association program was recently demonstrated by Geotech with an improved detection processor (DP), for April 1, 1979. Of 25 analyst generated events, our AA produced 22. One failure was due to missed weak signals, and another two were due to the large error in location resulting from small detection mistimings on PKP phases. After all available beam information has been purged from the signal arrival list, it is, of course, necessary for AA to use an algorithm like COMBO to attempt to associate the remaining unassociated signals.

AA preduced 49 false alarms from this day's dataset. This large number consists mostly of 3 station events from the Alaskan net or from COMBO.

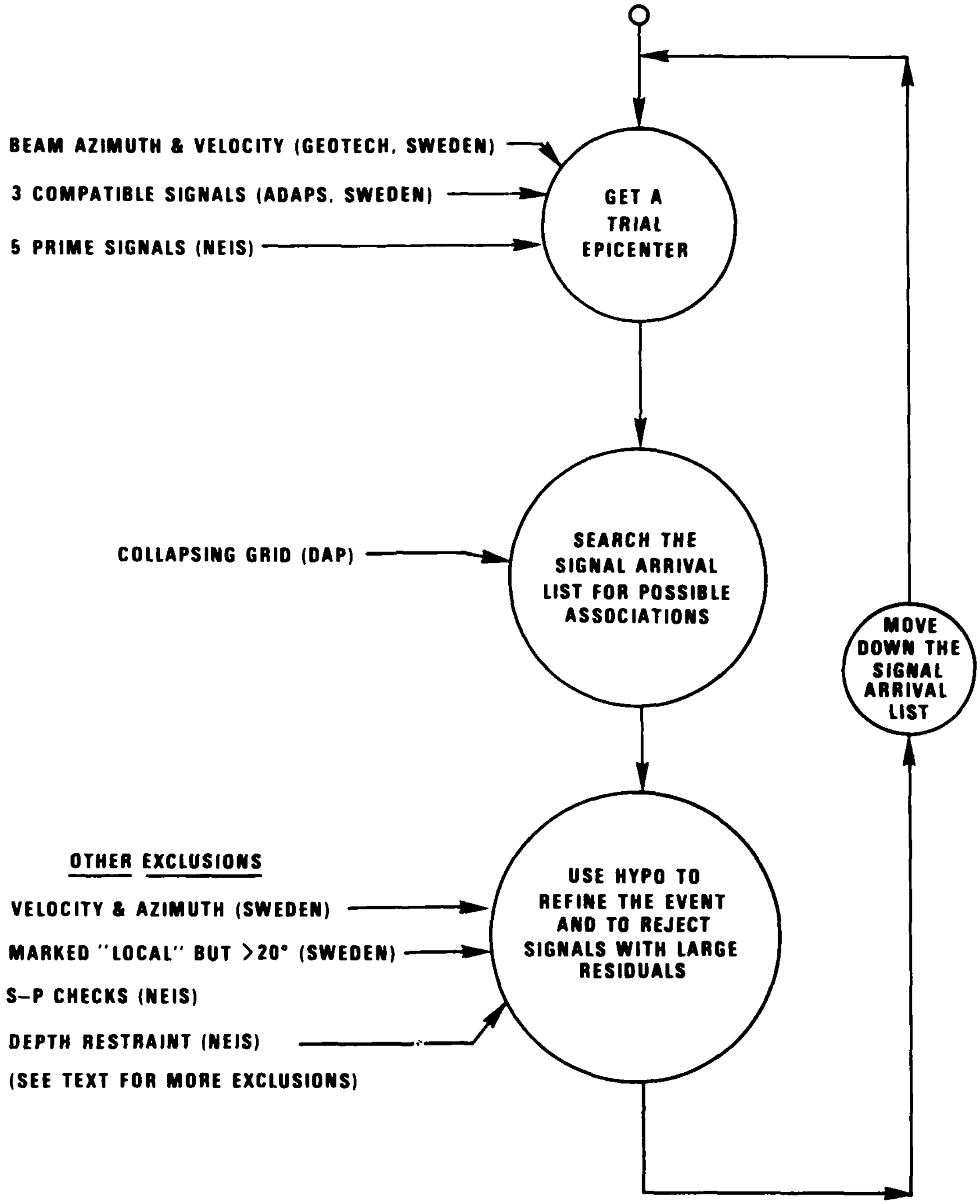

Figure 7. Generalized schema used by automatic seismological association computer programs.

Depth Considerations

An array detection is reported as a signal direction and velocity, from which the distance from station to epicenter can be determined by table look-up. AA does not use depth as a parameter in the lookup, but assumes zero depth of focus. An earthquake located on the ray path from this location back to the receiving station would produce a signal with the same azimuth and velocity; and if the real event is deep, its look-up location will be in error. The magnitude of the error increases with depth and closeness to the receiving station. The predicted arrival time at a station can have an error of 50 seconds for an event at a depth of 286 km and a distance of 25°, depending on the location of the stations.

This error in the SDAC AA shows up as an inability to pick out possible associatable arrivals from the signal queue, and has been observed when WRA was the driving station. The predicted arrival time from the trial location (plus the allowance window) is too far from the observed arrival time to catch the signal. Of course, the allowance window can be enlarged, and we found that increasing the allowance window from 100 sec to 200 sec has enabled arrivals to be found, so that events proceeded to successful generation. Obviously, though, it would be unwise to permit the allowance window to get too large, as this would simply collect unwanted signal arrivals which could distort or otherwise alter the event refinement in HYPO.

The point to be recognized here is that there is an interplay between the data, the algorithms and the parameters in this automatic association program. We believe that improving the performance of AA can be achieved by analyzing the quality of available data and the results of algorithms and parameters in order to derive a system which maximizes effective use of the available data.

DP process should be analyzed from the point of view of AA's needs for more effective data and DP should be changed accordingly, if practical.

Array beam processing should be similarly analyzed for ways of producing more effective data for AA.

It should not be forgotten that AA results of operating on "real-world" data should produce "real-world" results. At the present time, to our knowledge, the best gauge with which to measure AA's performance is the bulletin produced by an experienced senior analyst working the same dataset, but with waveform information at his hand, also.

RECOMMENDATIONS

The objective of this evaluation study was to examine the strength and weaknesses of state-of-the-art automatic association programs, and to see, in particular, if there existed algorithms and stategies which could be incorporated into Geotech's AA or ADAPS to improve the performance of these programs.

In Table II are listed in matrix form the major algorithms encountered in this study and the programs that uses them. This tabulation seems to show that ADAPS has almost all the algorithms known except as noted in the following paragraph; whereas, despite being a recent successor, Geotech's AA appears somewhat skeletal.

- It is recommended that three-way compatible station combinations (COMBO) be incorporated in the AA.
- It is recommended experimentation be done with the AA (ADAPS also, if possible) to determine if these programs would benefit from pP depth restraint, and from including depth as a parameter to find initial arrivals. The depth phase pP would come from analyst picks as there is no way at present that an automatic detector can find that phase. Once it is a signal attribute, it could be recognized by a suitable algorithm.
- Later phase association has long been recognized as needed in AA.
- The Swedish work on probabilistic event declaration ought to be followed closely. While not working satisfactorily yet (see Dr. Blandford's memo), it should help greatly to eliminate false alarms. It is recommended that similar work be initiated at the SDAC.
- Experiment with "local" characterization. The improved DP generates this signal attribute based on the presence of high frequencies. Should at least appear on printout, and may be used to exclude a signal if event is more than 20° distant.
- Similar to the Swedish program, check observed azimuth and slowness

TABLE II

Algorithms and Strategies Found in the Five Automatic Association Programs.

	AA	ADAPS	DEMO	FORAGE_S	DAP
Use beam azimuth and velocity to find trial epicenter	X		X		X
3 compatible signals to find trial epicenter (COMBO)		X	X		
5 prime signals to find trial epicenter				X	
Shrinking Grid, eliminates trial epicenter					X
S-P checks		X	X	X	
Use PKPDF for trial epi-center and for event refinement	X				
HYPO-type location program used iteratively to reject signals with large residuals	X	X	X	X	X
Depth restraint with pP				X	
Depth search is part of find trial epicenter			X		X
Seismologist interaction possible		X		X	
Associate to existing event when new data is added		X		X	
Special processing for after shock sequences		X			
Tripartite arrays to find trial epicenters	X	X			

TABLE II Continued

Algorithms and Strategies Found in the Five Automatic Association Programs.

	AA	ADAPS	DEMO	FORAGE_S	DAP
Use S-P phase to find trial epicenter		X			
Associate later phases		X			
Check: Is travel time between two stations greater than difference in arrival times?		X	X	X	
"Local" event excluded if distance > 20°		X	X		
Minimum number of signals needed to declare an event	(3)	(3)	(4)	(5)	(?)
Probabalistic approach for excluding stations (not working well, yet)			X		
Check: Do observed azimuth and slowness agree with that computed from event and station locations			X		

with values computed from station and event location. Exclude if residual is high. This can also be done if P-S phase distance is available.

- This developmental work will use the month of April 1979 for the data base.

- Plans should be undertaken for extending AA's capability to include regional distance as well as the present teleseismic. This will include P_n, P_g and L_g. The transfer should occur when another computer, probably VAX, is available on which the editing facility exists.

ACKNOWLEDGEMENTs

The author wishes to acknowledge Dr. Blandford's close interest in this subject, and also P. Kovacs' many years of work with ADAPS and AA. Dr. A. Chang provided tables to study the mislocation problem when zero depth is assumed. Brian Carroll had major responsibility in organizing the structure and in implementing the SDAC AA two years ago.

This work was monitored by the Air Force Technical Applications Center and was performed under contract # F33600-79-C-0549.

REFERENCES

Dahlman, O. and A. Sidor (1979). International seismological datacenter, demonstration facilities in Sweden, (preprint) *Försvarets Forskningsanstalt*, Stockholm.

Engdahl, E. R. and R. H. Gunst (1966). Use of a high speed computer for the preliminary determination of earthquake hypocenters, Bull. Seism. Soc. Am., 56, 325-336.

Snell, N. S. (1978). The detection association processor, Texas Instruments, Inc., Alexandria, Virginia, ALEX(01)-TR-77-06.